

REMARKS

In the office action mailed on August 12, 2004, claims 40–70 and 72–80 are rejected under 35 U.S.C. 103 over U.S. Patent No. 6,336,124 to Alam et al (“*Alam*”), and claim 71 is rejected over *Alam* in view of U.S. Patent No. 6,125,391 to Meltzer et al. (“*Meltzer*”). Applicant traverses the rejections and requests reconsideration in light of the following remarks.

Applicant and the undersigned thank Examiners Bashore and Nguyen for their time during the telephonic interview conducted on September 24, 2004.

The application describes, in part, the translation of a source document into an internal representation of the document. *Alam* describes, in part, the conversion of a source document into an intermediate format document. The internal representation, and the methods of translating documents into the internal representation disclosed in the application and recited in the application’s claims are patentably distinct from the intermediate format and conversion method described in *Alam*.

Alam Fails to Teach or Suggest the Elements of Claim 40 Asserted by the Action

Claims 40–70 and 72–80 are rejected under 35 U.S.C. 103 as being obvious in light of *Alam* and knowledge known to those skilled in the art. Claim 71 is rejected over *Alam* in view of *Meltzer*. The Action asserts that *Alam* teaches or suggests all elements of claims 40–70 and 72–80 other than the use of pointers that “associate the document objects within the document structure with the data content stored in the data content structure.” Applicant respectfully disagrees with this assertion.

Specifically with respect to independent claim 40, in addition to the failure to teach or suggest the recited use of pointers, *Alam* fails to teach or suggest at least the following limitations:

using the identified document agent to translate the source data into an internal representation of the digital content, the translation including:

for each object identified within the source data, creating a document object that represents an internal representation of the identified object and that separates the structure of the object from the data content of the object,

The above referenced language of claim 40 specifically indicates that in creating a document object, the structure of an object is separated from the data content of an object.

One non-limiting example of such a separation is described on page 13 of the specification at lines 12–16 as follows:.

“As can be seen from the above, the format of the internal representation 14 separates the ‘structure’ (or ‘layout’) of the documents, as described by the object types and their parameters, from the ‘content’ of the various objects; e.g. the character string (content) of a text object is separated from the dimensional parameters of the object; the image data (content) of a graphic object is separated from its dimensional parameters.”

The separation of structure and content includes the creation of two separate data structures that form part of the internal representation, with pointers providing references between the data in the structures. The creation of these data structures is explicitly recited in the independent claim 40, and one non-limiting description of is set forth in the specification on pages 16–17. *Alam* fails to teach or suggest the generation of such distinct data structures.

In stark contrast, in generating an intermediate format document, *Alam* interleaves structure and content. The Action relies upon column 2, lines 12–27 and column 7, lines 10–21 of *Alam* to suggest this limitation. The text in column 7, lines 10–21 relates to Figure 7. Figure 7 illustrates *Alam*’s method of converting data to an intermediate format. The steps illustrated in Figure 7 and described in the referenced text include:

- Locate and store tags, if any, in the input format document
- Locate all words in the input format document;
- *Join* words into lines;
- *Join* lines into paragraphs;
- Locate tables; and
- Output data in intermediate format.

(Emphasis Added) To the extent that this figure discloses structural objects within the intermediate format document, e.g., lines and paragraphs, the content, e.g., words associated with those structural objects, is explicitly *joined* with the structure itself. Column 2, lines 12–27 provide a similar description. The text states:

A computer implemented method of converting a document in an input format to a document in a different output format is disclosed. The method generally comprises locating data in the input document, *grouping data into one or more intermediate format blocks* in an intermediate format document, and converting the intermediate format document to the output format document using the intermediate format blocks. Preferably, the grouping includes *locating words* in the input document, *joining words satisfying line threshold to into lines*, *joining lines satisfying paragraph threshold into paragraphs*, and locating tables. The grouping may alternatively or further include locating tags (or control characters) in the input document and utilizing the tags in locating words, joining words into lines, joining lines into paragraph, and locating tables. Each intermediate format block may be selected from a word, a line, a paragraph, a table, and an image.

(emphasis added). This section also describes combining content with structure. Similar to the steps illustrated in Figure 7, the above quoted excerpt includes “grouping data into one or more intermediate format blocks”, “joining words satisfying line threshold into lines”, and “joining lines satisfying paragraph threshold into paragraphs”. Each of these steps suggests combining content with structure as opposed to separating content from structure as recited in claim 40.

Furthermore, even the most basic content elements that *Alam* deals with, words, are matched with structural information in *Alam*. For example:

In locating words from digital data representing an image of a document at step 702, the digital computer utilizes information provided for each word by the digital data in an input format. The information provided by the digital data in the original input format may include, for example, X and Y coordinates for the top left and bottom right of the word relative to the page as well as the font of the word. The font information includes information on the style, size, weight (bold or non-bold), stroke (italics or non-italics) and orientation of the word.... The individual words are then sorted by their X and Y coordinates, preferably first by the Y coordinate in the vertical direction and then by the X coordinate in the horizontal direction.

Col. 7, ln. 60 – col. 8, ln. 6. In contrast to the claims of the present application that recite separation of structure and content, according to above text, in *Alam*, structural information, specifically the X,Y coordinates of words within a document, forms the basis of content organization.

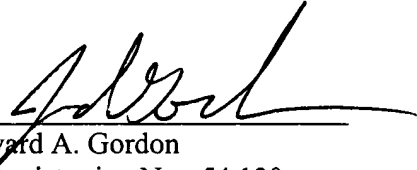
In light of the above, Applicant submits that *Alam* fails to teach or suggest all limitations of the translation element of claim 40. Applicant therefore requests reconsideration and withdrawal of the 103 rejection of claim 40. Claims 41–70 and 72–74 depend upon claim 40 and add further limitations thereto. Applicant therefore requests reconsideration and withdrawal of rejections of these claims, as well. Independent claims 75 and 76 include similar limitations as claim 40, and claims 77–80 add additional limitations to claim 76. Applicant therefore requests reconsideration and withdrawal of the rejections of claims 75–80.

Claim 71 is rejected under 35 U.S.C. 103 over *Alam* in view of *Meltzer*. *Meltzer* fails to cure the deficiencies of *Alam*. Therefore, the combination of *Alam* and *Meltzer* fails to teach all elements of claim 71. In view of the above, Applicant respectfully requests reconsideration and withdrawal of the 103 rejection of claim 71.

Applicant believes no fee is due with this response. However, if a fee is due, please charge our Deposit Account No. 18-1945, under Order No. PGLD-P01-004 from which the undersigned is authorized to draw.

Dated: October 13, 2004

Respectfully submitted,

By 
Edward A. Gordon
Registration No.: 54,130
ROPES & GRAY LLP
One International Place
Boston, Massachusetts 02110-2624
(617) 951-7000
(617) 951-7050 (Fax)
Attorneys/Agents For Applicant